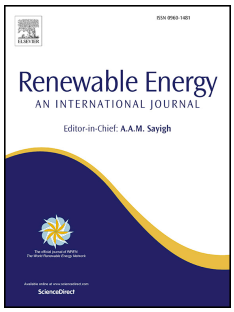
Accepted Manuscript



A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in China

Qingli Dong, Yuhuan Sun, Peizhi Li

PII: S0960-1481(16)30896-5

DOI: 10.1016/j.renene.2016.10.030

Reference: RENE 8218

To appear in: *Renewable Energy*

Received Date: 8 January 2016

Revised Date: 13 August 2016

Accepted Date: 16 October 2016

Please cite this article as: Dong Q, Sun Y, Li P, A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in China, *Renewable Energy* (2016), doi: 10.1016/j.renene.2016.10.030.

A novel forecasting model based on a hybrid processing strategy and an optimized local linear fuzzy neural network to make wind power forecasting: A case study of wind farms in China

Qingli Dong, Yuhuan Sun* and Peizhi Li
Dongbei University of Finance & Economics, Dalian, China

Abstract

As a crucial issue in the wind power industry, it is a tough and challenging task to predict the wind power accurately because of its nonlinearity, non-stationary and chaos. In this paper, we propose a novel hybrid model, which combines an integrated processing strategy and an optimized local linear fuzzy neural network, to forecast the wind power. First, discrete wavelet transform and singular spectrum analysis are used to filter out the noises and extract the trends from original wind power series, respectively. Then, the novel no-negative-constraint-combination theory together with the CS algorithm are adopted to integrate these two subseries obtained from the first step to retain strengths of each method. Based on the phase space reconstruction model, we could determine the most proper structure of the input sets and the output sets. Next, the local linear fuzzy neural network, with the initial rule consequent parameters optimized by the seeker optimization algorithm, is utilized to make wind power forecasts for a selected number of forward time steps. The numerical results from two experiments demonstrate that the proposed hybrid model is an effective approach to predict wind power, and the accuracy of prediction is highly improved compared with conventional forecasting models.

Keywords: Wind power; Processing strategy; Local linear fuzzy neural network; Forecasting

1. Introduction

Along with the rapid development of information science and the rapid integration of traditional industries and intelligent technology, big data analysis has penetrated into many industries, including renewable sources, storage, energy management and monitoring systems, which are the components of “smart grids” currently under deployment [1]. The operation of such a large and complex electric system, which depends on the coordination of all subsystems, can generate massive amounts of data. The challenges of data management and analysis are particularly important for the deployment of smart grids.

In fact, it is critically important to find an effective tool that can provide accurate wind power predictions, which could benefit the efficient management of the power grid and help mitigate the unfavorable effects in the growing wind energy scenario as well as increase the revenues from the electricity market with the optimization of bidding strategies [2]. Nevertheless, due to its high correlation with fluctuating and volatile wind speeds and many other meteorological conditions, wind energy is inherently intermittent, making it difficult to accurately forecast, which further gives rise to a relatively weak grid for large scale wind power input to the system [3]. Thus, inaccurate wind power forecasting may result in a series of problems.

In recent decades, different kinds of models have been developed to handle wind power prediction. Generally, there is a similar format when conducting wind power studies: first, applying an integration of meteorological models and historical wind plant output, while wind data about the research question are collected for a set of

potentially viable plant locations. Then, the historical data of wind speed and load, whose seasonality and other characteristics are correlated, are collected for the same period. Thereafter, preprocessed data are adopted in certain types of models, such as power system models, statistical analyses and so on, to assess the stock or provide accurate forecasts.

To be specific, the existing models, which have been proven effective in predicting wind power or speed, can be divided into three categories: (i) physical models, (ii) conventional statistical models and (iii) intelligent forecasting methods [4].

Physical models mainly include models based on spatial correlation, Numerical Weather Prediction (NWP) and so on. Power output are usually calculated by substituting them into the manufacture power curves, which utilize weather prediction data such as pressure, obstacles, temperature and surface roughness to solve complex mathematical models [5]. Salcedo-Sanz et al. [6] integrated two different meteorological prediction global models to forecast the mean hourly wind speed prediction. Cassola & Burlando [7] proposed a model that combines the NWP data and Kalman filtering to predict the wind energy. A new method, termed the CS-FS-WRF-E model, which integrates a Weather Research and Forecasting (WRF, one of the applications of NWP), a cuckoo search algorithm and a fuzzy system , was proposed by Zhao et al. [8]. Both the models yield improvements in the forecasting performance to a certain degree, however, there are some limitations when using approaches involving NWP models to perform long-term wind power forecasting because it exhibits satisfied forecasting performance only when the environment remains stable [9] and the errors of NWP forecasting are larger than those produced by many other types of models for short forecasting horizons (e.g., 3-6 h) [11,12]. Moreover, these methods are very complex and time-consuming, making them difficult to implement [12].

Conventional statistical models emulate the relationship between the historical values of wind power and the historical values of meteorological variables, and most of the parameters are estimated from data without making any assumptions regarding the physics and mathematics of the phenomenon under study [13]. The most representative models are the Autoregressive model, the Persistent Random Walk model (PRW) and the Autoregressive Integrated Moving Average (ARIMA) model [14]. Riahy & Abedi [15] proposed a model that utilized the linear prediction method in conjunction with filtering of the wind speed waveform to make wind speed forecasting. Besides, Autoregressive Moving Average process (ARMA) and persistence models were integrated by Torres et al. [16] to predict the hourly average wind speed 10h in advance. Moreover, Khashei et al. [14] integrated the neural networks and fuzzy logic with traditional ARIMA model to overcome the obstacles of ARIMA, including the linear data limitations, and this attempt exhibited obvious improvements in forecasting accuracy. Generally, these statistical models can outperform physical methods in short-term horizon forecasting, and they have been proven to be simple and effective approaches, although current trends are focused on intelligent forecasting methods.

Thanks to the rapid development of artificial intelligence technology, intelligent forecasting methods have been applied widely for wind power/speed prediction, among which ANNs (Artificial Neural Networks) [17], SVMs (Support Vector Machines) [18], Bayesian methods [19] and Fuzzy logic are discussed mostly. Li & Shi [20] presented a comprehensive comparison between three different ANNs in 1 h ahead wind speed prediction, including adaptive linear element, radial basis function

and back-propagation network. Several novel SVMs structures were proposed by Ortiz-García et al. [21] to manage the diversity in input data and several parameterizations of a mesoscale model, and this model performed better than a similar multilayer perceptron. Besides, Salcedo-Sanz et al. [22] introduced a scheme which combines elements from Coral Reefs Optimization (CRO) algorithm and Harmony Search (HS) approach with an Extreme Learning Machine (ELMs) network, and the main characteristics and performances were elaborated in case studies. In general, intelligence prediction models can outperform the time series models at most time scales because of the strong nonlinear prediction capacity. Nevertheless, the accuracy and performances largely depend on the quality and sensitivity of the historical datasets, thus it is necessary to preprocess the datasets, such as interpolation processing, decomposition, seasonal adjustments and so on [23]. Zhang et al. [13] utilized the Wavelet Transform technique to de-noise original series while Seasonal Adjustment Method (SAM) was used to separate the seasonal items, and final forecasting outputs were obtained by Radial Basis Function (RBF) network. Gnana Sheela & Deepa [24] proposed a hybrid computing model which integrates Multilayer Perceptron network and Self Organization Maps (SOM) to make wind speed forecasting. The aim of SOM is to maximize the degree of similarity of pattern within a cluster and then present the results in a lower dimensional space. The experimental results showed that the hybrid models have better performance than single methods. Liu et al. [25] developed a hybrid model based on Elman neural networks and the secondary decomposition algorithm, which combines the Wavelet Packet Decomposition and the Fast Ensemble Empirical Mode Decomposition, to predict the wind speed with high accuracy. Their results also validated the effectiveness of these hybrid models.

From the reviewing and comparisons, we can see that hybrid models are superior to single models in most cases, which is not hard to imagine as for hybrid or combination models, the basic idea is to combine different methods retaining strengths of each approaches to optimize the whole. However, most studies emphasize the processing of whole series, for instance the single Empirical Mode Decomposition (EMD), while neglecting the parallel structure of datasets. Based on the processing of historical wind power data, the short term wind power prediction is focused in this paper as follows: the Discrete Wavelet Transform (DWT) method is applied to filter out the noise signals from original datasets while Singular Spectrum Analysis (SSA) is presented as the tool for trend extraction. To combine DWT and SSA approaches retaining strengths of each method, herein the no negative constraint combination theory (NNCT) [26] is applied to integrate these two approaches, aiming at taking advantage of each characteristics. Next, the Phase Space Reconstruction (PSR) method is adopted to select the most proper parallel structure of the input sets, before we present the optimized local linear fuzzy neural network as the predictor to make short term wind power forecasting.

The main contributions of this study are given as follows:

i. As a critical procedure of time series analysis, it is difficult to identify each component as trend and quasi-periodic components while extracting the noise parts. A key advantage of discrete wavelet transform over Fourier transforms is temporal resolution: it captures both frequency and location information. Besides, singular spectrum analysis is a nonparametric spectral estimation method, which combines elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing, and it can extract amplitude-modulated sine wave components [27]. Herein, it is the first try

to integrate them to process wind power time records to improve the performance of forecasting.

ii. To select the proper structure of input set and output set is also a tough question, which can affect the mapping results a lot. Furthermore, the forecasting accuracy depends on whether previous sequential data can reveal the dynamical properties of the wind power system, and how sensitive are the chosen input sets. In this paper, with the delay time and the embedding dimension determined by the mutual information method and G-P method, respectively, phase space reconstruction model is applied to select the proper form of dataset and the experiments results verify its effectiveness.

iii. Trained by a locally linear model tree learning algorithm, the local linear fuzzy neural network has been validated to be effective and has good generalization performance in forecasting fields. In this paper, the SOA is adopted to optimize the initial rule consequent parameters to improve convergence speed and reduce calculation amount. Optimized network has further advantages with characteristics of higher reliability and avoid of over-fitting. Finally, two numerical modeling experiments conducted in different condition verify the superiority of the hybrid model, which has obvious improvements in the accuracy of wind power forecasting.

The remainder of this paper are organized as follows. In section 2, we briefly discuss the processing strategy. Then, the proposed forecasting model will be discussed and analyzed in section 3. In section 4, the forecasting performance evaluation methods are introduced. The simulation and numerical results of our hybrid model, and the results of the comprehensive comparison between different models are discussed in section 5. Finally, section 6 provides the conclusions of this study.

2. Processing strategy

The development of datasets processing can be divided into four phases. First, the raw wind power data obtained from wind power plants may contain missing values inevitably, it will be addressed by cubic spline interpolation [28] method. Second, DWT and SSA are conducted to process original datasets, respectively. Then, the NNCT together with the CS algorithm will be applied to integrate these two subseries obtained before. Finally, we adopt PSR approach to choose the proper structure of input and output sets. The flow chart of the developed hybrid processing strategy is presented in Fig. 1.

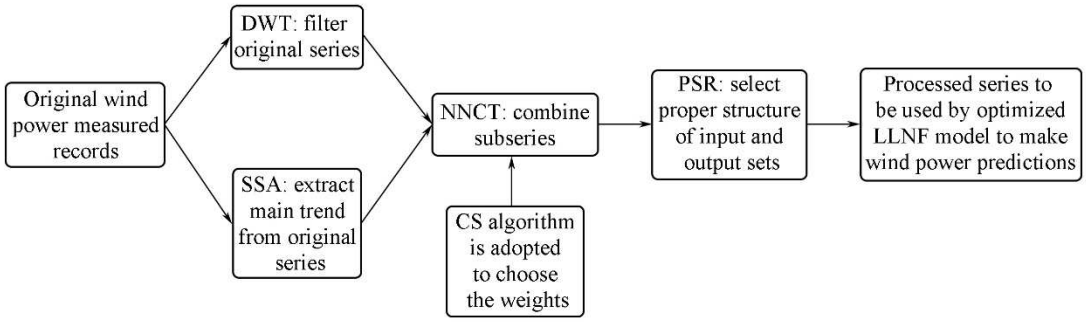


Fig. 1. Framework of the developed hybrid processing strategy.

2.1. The discrete wavelet transform

Due to the rate-change operators in the filter bank, the DWT is not time-invariant

but actually very sensitive to the alignment of the signal in time. To address the time-varying problem of wavelet transforms, Ref. [29] proposed a new algorithm for wavelet representation of a signal, which is invariant to time shifts. According to this algorithm, only the scale parameter is sampled along the dyadic sequence and the wavelet transform is calculated for each point in time. The wavelet analysis results indicate the comparability between the original signal and particular wavelet. A proper wavelet basis function can affect the performance of the signal filtering considerably. The corresponding wavelet family has many daughter wavelets, which are expressed as follows:

$$\boldsymbol{\psi}_{x,y}(t) = a_o^{-x/2}\boldsymbol{\psi}\left(\frac{t - b_0 a_0^x}{a_0^x}\right) \tag{1}$$

where integers x and y control the wavelet dilation and translation, respectively. $a_0 \in \mathbb{Z}$, $b_0 \in \mathbb{Z}$. While a_0 is the scale factor and b_0 denotes the shift factor, whose function is to maintain energy preservation.

Definition: The dyadic wavelet can be described as:

$$\boldsymbol{\psi}_{x,y}(t) = 2^{-x/2}\boldsymbol{\psi}\left(2^{-x}\cdot t - y\right) \tag{2}$$

where the coefficient $2^{-x/2}$ is a normalization constant and $L^2(R)$ is the norm for energy normalization of wavelet basis. The DWT is proposed to be the inner product of the original signal and the mother wavelet function; thus, the wavelet coefficient of DWT can be deduced as:

$$\boldsymbol{W}_f(x,y) = \left\langle f(t), \boldsymbol{\psi}_{x,y}(t)\right\rangle = 2^{-x/2}\int f(t)\boldsymbol{\psi}\left(2^{-x}\cdot t - y\right)dt \tag{3}$$

where $\boldsymbol{\psi}$ denotes wavelet basis and $f \in L^2(R)$:

$$f(t) = \sum_x\sum_y \boldsymbol{W}_f\left(2^{-x}, 2^{-x}\cdot y\right)\boldsymbol{\psi}\left(2^{-x}\cdot t - y\right) \tag{4}$$

Remark: The above process of the DWT can be regarded as a filter process, which passes the band-pass filter, the DWT and its reconstructed version may be described in a dyadic form [30]. Thus, during the DWT process, it is feasible to select a periodic function for boundary treatment to ensure a favorable effect.

2.2. Singular spectrum analysis

As a relatively new method, the SSA method is more advantageous than classical time series techniques for time series analysis and forecasting. Based on SVD, the SSA technique is capable of identifying the original series as several independent components, including the trend, periodic, oscillations, noise and 'clean' series. According to Hassani et al. [31], the SSA algorithm has two main stages: decomposing and reconstruction. During the decomposing process, the first step is concerned with mapping a one-dimensional series into the multi-dimensional series, then the singular value decomposition of the trajectory matrix is obtained. In reconstruction stage, the elementary matrices are split into several groups, then we perform diagonal averaging to transform each matrix into a time series. For more information, see Ref [27]. The algorithm is described as follows.

Algorithm 1: singular spectrum analysis

Input: $X = \left(x_1, x_2, \ldots, x_N\right)^T$ –a sequence of raw data.

Output: $Y = \left(y_1, y_2, \ldots, y_N\right)^T$

Parameters: N -the length of vector.
L -the window length.

/* Form the trajectory matrix into a Hankel matrix */
FOR (int $i=1;\ i<=N-L+1;\ ++i$)
$Z_i=(x_i,x_{i+1},\ldots,x_{i-L+1})^T$;
END $Z=[Z_1,Z_2,\ldots,Z_{N-L+1}]$;
/* Perform the singular value decomposition of the trajectory matrix */
$d=\max\{i\,|\,eigenvalues(ZZ^T)>0\}$;
/* $U_i(i=1,\ldots,L)$ are the orthonormal system of the eigenvectors */
FOR (int $i=1;\ i<=d;\ ++i$)
$\sqrt{\lambda_i}V_i=Z^TU_i$;
$Z_i=\sqrt{\lambda_i}U_iV_i^T$;
END $Z=[Z_1,Z_2,\ldots,Z_d]$
/* Partition the set of indices $\{1,\ldots,d\}$ into m disjoint subsets $I_1,\ldots,I_m$ */
$Z=Z_{I_1}+Z_{I_2}+\ldots+Z_{I_m}$;
/* Each matrix Z_{I_j} is Hankelized and then the obtained Hankel matrix is transformed into a new series with the length of N */
/* Diagonal averaging applying to a resultant matrix $Z=(z_{ij})_{L\times K}$ produces a reconstructed series $Y=(y_1,y_2,\ldots,y_N)^T$ */
IF $L<K$ $z_{ij}^*=z_{ij}$; **ELSE** $z_{ij}^*=z_{ji}$; **END**
FOR (int $n=1;\ n<=\min(L,K);\ ++n$)
$y_n=mean(z_{1,n+1}^*,z_{2,n}^*,\ldots,z_{n+1,1}^*)$;
END FOR (int $n=\min(L,K);\ n<=\max(L,K);\ ++n$)
$y_n=mean(z_{1,n+1}^*,z_{2,n}^*,\ldots,z_{\min(L,K),n-\min(L,K)+2}^*)$
END FOR (int $n=\max(L,K);\ n<=N;\ ++n$)
$y_n=mean(z_{n-\max(L,K)+1,\max(L,K)}^*,\ldots,z_{N-\max(L,K)+1,\max(L,K)+n-N}^*)$
END
RETURN Y

2.3. No negative constraint combination theory

As is mentioned above, DWT and SSA can address the raw signal from different aspects. To be specific, the former can extract diagnostic information from non-stationary series, because it can not only filter out the white noise or useless information presented with the selected predictors, but preserve the dyadic properties of the decomposed signal. The latter is often applied to identify and extract periodic, quasi-periodic and oscillatory components from the original signal.

As a widely accepted method to improve the performance of individual approaches, the traditional combination theory attempts to find the best weight based on minimizing the sum of squared errors:

$$\min \Re = \boldsymbol{L}^T \boldsymbol{E} \boldsymbol{L} = \sum_{k=1}^{K} \sum_{j=1}^{m} \sum_{i=1}^{m} l_i l_j e_{ik}$$
$$st \begin{cases} \boldsymbol{R}^T \boldsymbol{L} = 1 \\ \boldsymbol{L} \geq 0 \end{cases} \tag{5}$$

where $\boldsymbol{L} = (l_1, \ldots, l_m)^T$ denotes the weight vector, $\boldsymbol{R} = (1, \ldots, 1)^T$ is a column vector and $\boldsymbol{E}_{ij} = e_i^T e_j$ is the element of combination error matrix.

However, there is no convincing proof and reason for the combination weights to be positive, on the contrary, the performance of combination methods with negative weights can improve, especially when the trend of individual performance are largely different or similar [26].

Based on the proposed NNCT, generalized combination is given as follows:

$$\min \Re = \boldsymbol{L}^T \boldsymbol{E} \boldsymbol{L} = \sum_{k=1}^{K} \sum_{j=1}^{m} \sum_{i=1}^{m} l_i l_j e_{ik}$$
$$st \ \boldsymbol{R}^T \boldsymbol{L} = 1 \tag{6}$$

This equation shows that there have no limitation of the weight vector in the range of $[0,1]$. This part supplies a novel weight-determined method which was testified by previous experimental simulation, and experiments in this paper can also verify the advantages to loosen the restrictions of weights.

2.4. Phase space reconstruction

PSR is a powerful tool which is widely used to search for patterns in a time series and in a higher dimension transformation of the time series. As the basis for chaotic time series analysis, Takens [32] proposed the delay-coordinate method to implement the PSR for the given time series $x = \{x(t) | t = 1, \ldots, N\}$:

$$\boldsymbol{X} = \left\{ \vec{X}_p \middle| \vec{X}_p = x_p + x_{p+\tau} + \ldots + x_{p+(m-1)\tau} \right\} \tag{7}$$

where $\tau \in \mathbb{N}^*$, $p = 1, \ldots, M$ and the phase space vector $\boldsymbol{X}$ is the matrix obtained by PSR, m and τ denote the embedding dimension and delay time respectively, and $M = N - (m-1)\tau$ is the points in phase space.

During PSR, the determination of m and τ has a lot with the accuracy of chaos forecast results. Theoretically, for time series without noise signal, there is barely limit for the delay time τ and embedding dimension m. This is also the reason to conduct the former process. In this paper, we apply the mutual information method to determine the delay time τ, and G-P method to determine the embedding dimension m. The algorithm description is shown as follows:

Algorithm 2: determine the best parameters of PSR

Input: $X(p)$ –matrix obtained by PSR.

Output: τ^* -Best value of delay time.
m^* -Best value of embedding dimension.

/* Due to the mutual information function measures the overall dependency, it is feasible for applied on wind power series */
FOR (int $t = 1$; $t <= M$; $++t$)

$I(t)=\sum p(x(t),x(t+\tau))\log\left[\left(p(x(t),x(t+\tau))\right)/\left(p(x(t))P(x(t+\tau))\right)\right];$

END

/* $p(\cdot)$ is the probability density function */

$\tau^{*}=\arg\min_{t\in\{1,\ldots,M\}} I(t);$

RETURN τ^{*}

/* Herein, G-P method proposed by Grassberge and Procaccia are employed to determine the embedding dimension m and delay time τ */

DO

/* Calculate correlation integral with positive r */

$C_n(r)=\lim_{n\to\infty}\frac{1}{M}\sum_i\sum_j\theta\left(r-\left\|X(i)-X(j)\right\|\right);$

/* $\theta(\cdot)$ is Heaviside of unit function */

$\theta(x)=\begin{cases}0, x\le 0\\ 1, x>0\end{cases};$

/* Select the appropriate r and obtain the similar structure of correlation dimension */

$D_m=\log C_n(r)/\log r;$

$m++;$

WHILE $(D_m-D_{m-1}\to 0)$

RETURN $m^{*}=\left\{m\left|D_m-D_{m-1}\to 0\right.\right\}$

3. Proposed forecasting model

In this section, the original LLNF and optimized LLNF will be discussed, as well as the optimization algorithm, termed the SOA. Besides, the hybrid model proposed in this paper will be introduced in detail.

3.1. Local linear fuzzy neural network

Based on the incremental tree learning algorithm, the LLNF utilizes a divide-and-conquer approach to solve complex problems by dividing the input space into small subspaces, each identified by a local linear model (*LLM*) and fuzzy validity function [27]. The complexity of the model increases with further improvements. There is one hidden layer and a linear neuron of the output layer in the overall fuzzy neural network, as illustrated in Fig. 2 (a).

The overall output of the LLNF model is as follows:

$$\hat{y}=\sum_{i=1}^{M}\hat{y}_i\phi_i(u) \tag{8}$$

where

$$\hat{y}_i=\eta_{i0}+\eta_{i1}u_1+\eta_{i2}u_2+\cdots+\eta_{ip}u_p \tag{9}$$

where $u=[u_1,\ldots,u_p]^T$ denotes the input value, M is the number of neurons, and $\hat{y}_i$ is the local output of each *LLM*, whose parameters are described as $\eta_i=\left[\eta_{i0},\ldots,\eta_{ip}\right]$.

To maintain the partition unity, the axis-orthogonal normal Gaussian function is adopted here as the validity function, which is expressed below.

$$\phi_i(u) = \mu_i(u) \Big/ \sum_{j=1}^{M} \mu_j(u), i = 1, \ldots, M \tag{10}$$

$$\mu_i(u) = \exp\left(-\frac{1}{2}\left((u_1 - c_{i1})^2 / \sigma_{i1}^2 + \cdots + (u_p - c_{ip})^2 / \sigma_{ip}^2 \right) \right) \tag{11}$$

where c_{ij} is the center coordinates and σ_{ij} is the standard deviation of the normalized Gaussian validity function.

From the transparent representation obtained by the LLNF model, the nonlinear hidden layer of the LLNF should be adjusted with a learning method. Experiments demonstrate that the initial rule consequent parameters η_{ij} can affect the accuracy of the model; therefore, in this paper, the SOA was adopted to determine the optimal initial parameters, and c_{ij} and σ_{ij} are modified by the locally linear model tree algorithm (LoLiMoT), which was proposed by Nelles et al. [33] and was based on the strategy of divide and conquer. This algorithm is different from others in its that it divides the input space by axis-orthogonal splits.

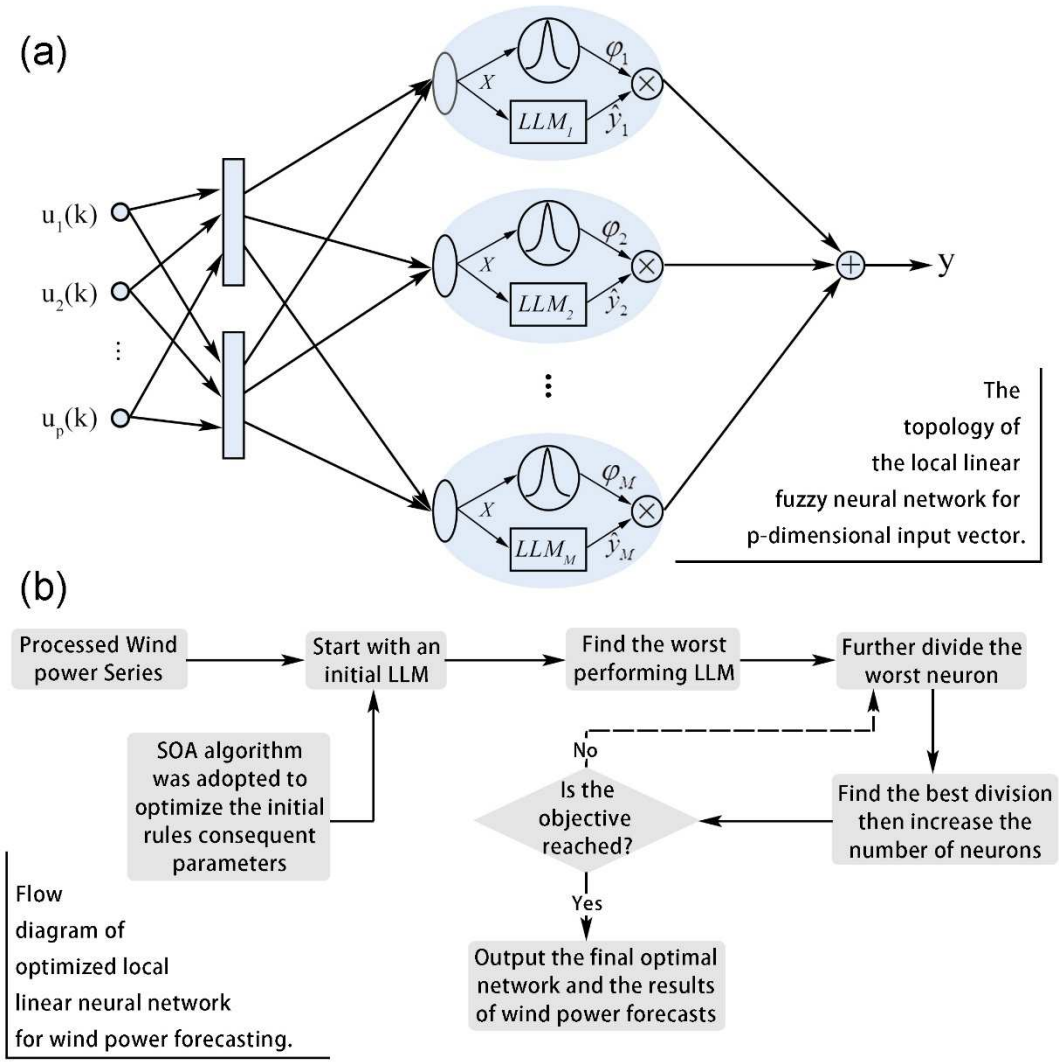


Fig. 2. The topology and flow chart of the optimal LLNF model.

3.2. Seeker optimization algorithm

Mimicking the human search process, the original SOA operates on a series of

solutions and individuals, which are called the human search team and seekers, respectively [34]. From social learning and cognitive learning, the SOA can conduct social experiments and cognitive experiments, respectively. Combined with the egotistic behaviour, altruistic behaviour and pro-activeness behaviour of the intelligent swarm, the searching direction of individuals can be determined.

Algorithm 3: seeker optimization algorithm

Input: $x_s = (x(1), x(2), \ldots, x(q))$ -a sequence of training data.

$x_p = (x(q+1), x(q+2), \ldots, x(q+d))$ -a sequence of verification data.

Generate N random seekers in the population;

Output: $\vec{\alpha}(t)$ -the value of α with the best fitness value.

Parameters:

s -the number of positions distribute uniformly and randomly in the search space.

$t = 0$ -the number of the initial iteration.

$\boldsymbol{\mu}_{ij}$ -the value of the membership function.

$\boldsymbol{\delta}_{ij}$ -the j th element of the vector.

/*Evaluate all seekers and save the historical best position.*/

DO

FOR ($int\ i = 1;\ i <= s;\ ++i$)

$\vec{\boldsymbol{d}}_{i,ego} = \vec{\boldsymbol{p}}_{i,best} - \vec{\boldsymbol{x}}_i(t)$;

$\vec{\boldsymbol{d}}_{i,alt} = \vec{\boldsymbol{g}}_{best} - \vec{\boldsymbol{x}}_i(t)$;

$\vec{\boldsymbol{d}}_{i,pro} = \vec{\boldsymbol{x}}_i(t_1) - \vec{\boldsymbol{x}}_i(t_2)$;

$\vec{\boldsymbol{\alpha}}_i(t) = \boldsymbol{\delta}_i \sqrt{-\ln(rand(\boldsymbol{\mu}_i))}$;

/* Update the i th seeker's position */

$\boldsymbol{x}_{ij}(t+1) = \boldsymbol{x}_{ij}(t) + \boldsymbol{\alpha}_{ij}(t) \cdot \boldsymbol{d}_{ij}(t)$;

$i++$;

END FOR

/*Evaluate all seekers and save the historical best position.*/

/*Implement the inter-subpopulation learning operation.*/

$t++$;

/*Until the termination criterion is satisfied.*/

END WHILE

RETURN $\vec{\alpha}(t)$

3.3. The proposed hybrid model

In this section, the proposed hybrid wind power forecasting model will be discussed in detail. The entire model can be divided into three parts:

Part one: Extraction and filtering.

It is known that disturbance and noise from data sources and signal transduction pathway are inevitable. DWT method, as is mention above, can filter the raw series of wind power, effectively. Furthermore, SSA approach is applied to extract periodic, quasi-periodic and oscillatory components from the primal data. Thus, two subseries of wind power can be obtained from individual methods, respectively.

Part two: Reconstruction and combination.

With the two processed series obtained in part one, the NNCT together with CS algorithm are applied to allocate proper general weights to each of them. Herein, the cuckoo search algorithm (CS) [8] is employed to select the best weights of each series for its faster convergence behaviour by compared with other common optimization algorithm. Thereafter, based on the PSR, the integrated wind power series from NNCT optimized by CS algorithm can be reconstructed into a proper matrix, accordingly, which is the input sets of the forecasting model.

Part three: Parameters optimization and forecasting

The LLNF divides the input space into small subspaces with fuzzy validity functions. The whole model is a fuzzy neural network, which has one hidden layer and a linear neuron in the output layer. During the process of LLNF, the initial rule consequent parameters are fine-tuned by SOA and the parameters of validity functions are reconciled through the LoLiMoT algorithm. After the model find the optimal number of neurons in iterations, the weighted sum of the outputs of *LLM* is computed as the final output. More details of this step are presented in Fig. 2 (b), which also are shown as follows:

Step 1: Start with an initial single *LLM*, whose rule consequent parameters are determined by applying the SOA over the entire input space.

Step 2: Find the worst-performing *LLM*, which has the maximum local loss function, e.g., the root-mean-squares of errors among all *LLMs*.

Step 3: The worst model (neuron) obtained in the former step, is considered for further division. The validation hypercube of this neuron is broken into two halves with an axis-orthogonal split. Divisions in all dimensions were carried out, and the following four steps are implemented for each of the divisions:

Step 3.1: Construction of the multi-dimensional fuzzy membership functions for both hyper-rectangles.

Step 3.2: Construction of all validity functions.

Step 3.3: Local estimation of the rule consequent parameters for both newly generated local linear neurons.

Step 3.4: Calculation of the global loss function for the current overall model.

Step 4: Find the best division, which is the best alternative checked in step 2, and increase the number of neurons.

Step 5: Inspect the convergence to determine whether the objectives are met.

Remark 1: Various options exist for the termination criterion, including a maximal model complexity, i.e., a maximal number of *LLMs*, statistical validation tests, or information criteria.

Remark 2: For the convenience of narrative, the proposed hybrid model is named as the DSNP-oLLNF model. (Considering that the DWT, SSA, NNCT, PSR and the optimized LLNF are included in our hybrid model).

4. Forecasting performance evaluation strategy

As there is no universal criterion to evaluate models, several common models quality metrics are adopted here. Moreover, the Diebold–Mariano (*DM*) test is an effective tool to compare the forecast accuracy of two forecast methods, which is also adopted in this paper.

4.1. Error measures

It is critical to select reasonable criteria to evaluate the final results of the experiments. Herein, the evaluation indicators are the root mean squares of the errors (*RMSE*) and standardized *RMSE*. In the same manner, the traditional criterion of the

mean-absolute of errors is also normalized as *NMAE*. The aforementioned indicators are frequently applied in the literature. The *NMAE*, *RMSE* and *sRMSE* both have normalized forms compared with the classic *MAE* and *MSE*, and these forms can describe and illustrate the prediction error in terms of MW power, which is typically of a high order of magnitude. The well-known mean absolute percentage error (*MAPE*) criteria is not considered because the actual value of wind power output may be zero at some points, which may lead to calculation errors. Moreover, there is also a common indicator to evaluate the wind power. The index of agreement (*IA*) is a normalized measure of the degree of model forecasting error, varying between 0 and 1. A value of 0 corresponds to no agreement, whereas 1 corresponds to a perfect match. This indicator can illustrate the proportional differences between the observed and forecasted means and variances. Both of these indicators are depicted in Table 1.

Table 1
Four error criteria.

Metric	Definition	Equation
RMSE	The root mean squares of the errors	$RMSE=\sqrt{\frac{1}{T}\sum_{i=1}^{T}\left(\boldsymbol{Y}_{ACT(i)}-\boldsymbol{Y}_{PRE(i)}\right)^2}$
sRMSE	The standardized root mean squares of the errors	$sRMSE=\sqrt{\frac{1}{T}\sum_{i=1}^{T}\left(\frac{\boldsymbol{Y}_{ACT(i)}-\boldsymbol{Y}_{PRE(i)}}{\boldsymbol{Y}_C}\right)^2}$
NMAE	The normalized mean-absolute of errors is also	$NMAE=\frac{1}{T}\sum_{i=1}^{T}\frac{\left\lvert\boldsymbol{Y}_{ACT(i)}-\boldsymbol{Y}_{PRE(i)}\right\rvert}{\boldsymbol{Y}_C}$
IA	The index of agreement	$IA=1-\frac{\sum_{i=1}^{T}\left(\boldsymbol{Y}_{PRE(i)}-\boldsymbol{Y}_{ACT(i)}\right)^2}{\sum_{i}^{T}\left(\left\lvert\boldsymbol{Y}_{PRE(i)}-\bar{\boldsymbol{Y}}_{ACT}\right\rvert+\left\lvert\boldsymbol{Y}_{ACT(i)}-\bar{\boldsymbol{Y}}_{ACT}\right\rvert\right)^2}$

Where $\boldsymbol{Y}_{PRE(i)}$ and $\boldsymbol{Y}_{ACT(i)}$ are the predicted and actual values of wind power at time point i of the forecasting horizon T, respectively. The capacity of the wind farm is denoted as $\boldsymbol{Y}_C$ and $\bar{\boldsymbol{Y}}_{ACT}$ is the mean of the actual values of wind power.

4.2. DM test

To maintain the integrity of the experiments logic, it is necessary to perform several statistical tests on our experiments. Herein, the *DM* test [35], which is widely used to evaluate predictive accuracy, is used to evaluate the performance of our hybrid DSNP-oLLNF model for wind power forecasting.

The *DM* test compares the forecast accuracy of two forecast models. Let $\boldsymbol{e}_1$ be the forecast error from method 1, whereas $\boldsymbol{e}_2$ is the forecast error from method 2. There is a loss function to measure the accuracy of each forecast method:

$$\boldsymbol{l}(\boldsymbol{e}_k)=\boldsymbol{e}_k^2,\{k=1,2,\ldots,n\} \tag{12}$$

Then, the loss differential between these methods can be defined as:

$$\boldsymbol{e}(h)=\boldsymbol{l}(\boldsymbol{e}_1)-\boldsymbol{l}(\boldsymbol{e}_2) \tag{13}$$

To determine whether model 1 has similar predict accuracy with model 2, the null hypothesis of the test can be defined as:

$$H_0 : \boldsymbol{E}\left[\boldsymbol{e}(h)\right] = 0 \tag{14}$$

Against the alternative hypothesis:

$$H_1 : \boldsymbol{E}\left[\boldsymbol{e}(h)\right] \neq 0 \tag{15}$$

The *DM* test statistic is computed as follows:

$$DM = \frac{\overline{\boldsymbol{e}}}{\left(\boldsymbol{S}^2/n\right)^{1/2}} \tag{16}$$

where $\overline{\boldsymbol{e}}$ is the sample mean of the loss differential and $\boldsymbol{S}^2$ is the consistent estimate of the variances of $\boldsymbol{e}(h)$.

The null hypothesis means that the two forecasting results have no notable differences, whereas the alternative hypothesis is that the accuracy of these two models are different. Therefore, the null hypothesis will be rejected at the α significance level once the *DM* statistic falls outside the confidence interval.

5. Experiments and discussion

Here, two experiments are conducted to validate the effectiveness of the proposed hybrid model and the improvement in accuracy for wind power forecasting compared with existing methods, using two databases from different locations in China. The geographic information and statistical description of datasets are listed in Table 2.

Table 2
Descriptive statistics of original wind power series involved in two experiments.

Data sources	Location	Data sets	Numbers	Means (KW)	Std. (KW)	Max (KW)	Min (KW)
Aksay	39.5° N, 94.5° E	All samples	38015	788.4119	336.8812	2077.14	0
		Training	29279	785.1507	335.2147	1948.75	0
		testing	8736	799.344	342.1805	2077.14	0
Yumen	40.5° N, 96.7° E	All samples	38682	402.4787	266.099	1936.13	0
		Training	29752	399.1539	271.4717	1936.13	0
		testing	8930	413.5573	247.0473	1293.4	0

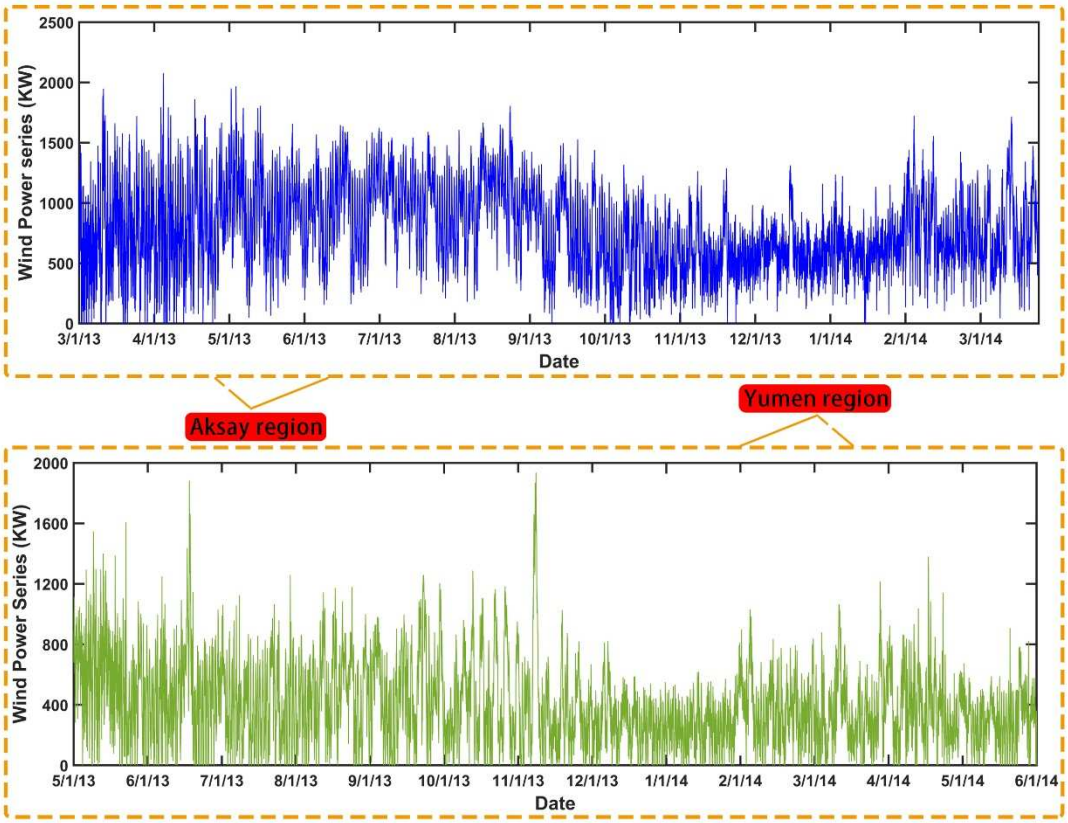


Fig. 3. Original wind power series from two data sources, with time interval of 15 minutes.

5.1. Experiment I

The Aksay region, which is located in Gansu Province, China, possesses abundant wind resources because of its unique geographical advantage and climatic superiority. The wind power data used in experiment I are collected from the Jinshan wind farm located in this region. Specific information regarding the farms' geographical position is provided in Table. 3.

There are numerous wind turbines in the wind farm, including typical and conventional turbines; the typical wind turbines can output an unlimited amount of wind power, and will be the focus of future studies. In this paper, all of the data series utilized are collected from these typical wind turbines, spanning from March 2013 to March 2014 with a time interval of **15 minutes**. Fig. 3 shows the pictures of 15 min average wind power from Aksay region along with the time variation. The fraction of the missing data is 2.03%, specifically, the number of entry/missing data is 38015/769.

As is introduced before, PSR method is adopted to reconstruct the processed power series into the most proper form as input sets and output sets. The critical procedure of PSR is the determination of the delay time τ and the embedding dimension m. In this paper, the G-D algorithm is utilized to select the τ and m, and the final values are determined as 2 and 4, respectively. According to the theory of PSR expressed as Eq. 7, we can obtain a single output equation, whose independent variables are historical data known:

$$x_{t+1+(m-1)\tau} = f\left(x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\right) \tag{17}$$

where $f(\cdot)$ is a mapping function which is the essence of forecasting. $x_t(t=1,2,\ldots,N)$ denotes the processed wind power series. When applying these

wind power series into models, we place a moving window on the time series to identify the training sets. In this experiment, the length of the training set covering an entire month. The form of the dataset used in the DSNP-oLLNF model, including the training set and testing set, are shown in Fig. 4. For instance, the data series collected from March 1 to March 31, 2013 is used to train the model, and then, the trained model is used to predict the following wind power series values of the first week in April. Thus, wind power data covering an entire year from typical wind turbines are utilized. The data horizon covers four seasons and all rapid fluctuations; these efforts are aimed at eliminating accidental errors and system errors of this experiment.

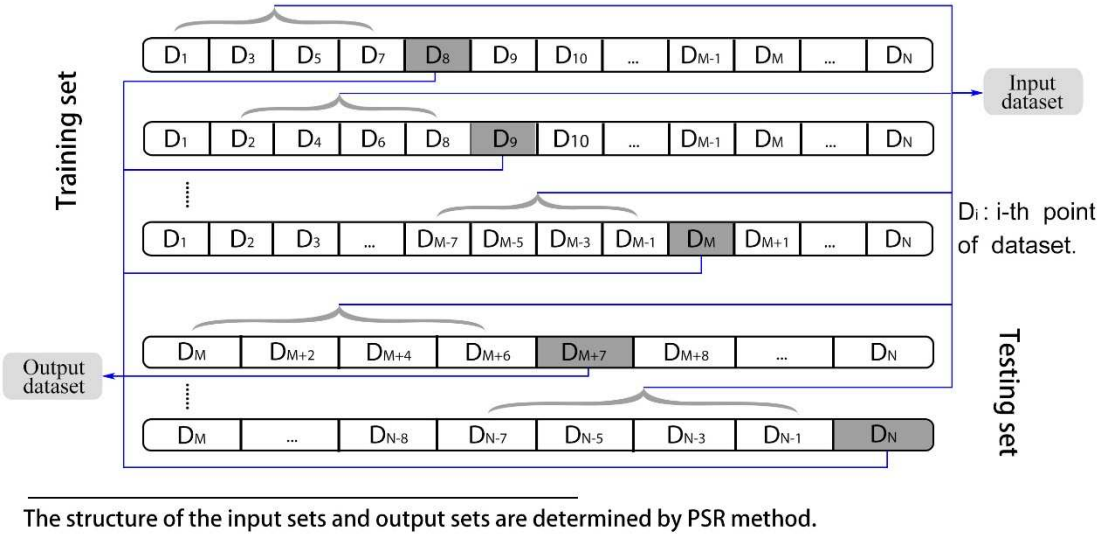


Fig. 4. Form of the dataset adopted in experiment I.

To ensure the fairness of the comparison, all methods considered in this experiment are evaluated by the same error criteria, namely, the *RMSE*, *sRMSE*, *NMAE* and *IA*. The comparison results obtained from experiments I are listed in Table 3, where the best performances have been identified in bold. The data shown in Table 3 are the forecasting results of the first week in April 2013, whereas the training sets are the data for the entire month of March 2013. To clearly show the comparison between different models, the forecasting results are divided into daily records. On April 1, the *RMSE* of the hybrid model (DSNP-oLLNF) is 0.0920, as computed by the forecasting results and test values of wind power. This error is smaller than those of the BPNN, ELMAN, GRNN and single LLNF models. The proposed method also has a smaller *RMSE* on April 2, 5, 6 and 7. On April 3 and 4, the *RMSE* value of the GRNN model is the smallest among all the reference models, which reflects the good performance of the GRNN model. Meanwhile, the mean value of the *RMSE* of the hybrid model is the smallest, representing an improvement of 14.2512%, 18.6246%, 12.6691% and 14.3546% over the BPNN, ELMAN, GRNN and single LLNF models, respectively.

The *sRMSE* is the standardized version of the *RMSE*; this parameter can neutralize the extreme values while reflecting the real differences. The *sRMSE* value of the DSNP-oLLNF model is 0.0580 on April 1, which is the smallest value of the models considered. The average *sRMSE* value of the hybrid model is 0.0909, presenting improvements of 17.663%, 21.84%, 16.1439% and 17.7376% over the BPNN, ELMAN, GRNN and single LLNF models, respectively. Moreover, the *NMAE*, which evaluates the performance from another perspective, also indicates the superiority of the DSNP-oLLNF model. Furthermore, the *IA* is used here to confirm

the validity of the overall performance. The average *IA* of the hybrid model is the highest among the models considered. In general, the merits of the hybrid model are clear based on all of the error indicators considered, including the *RMSE, sRMSE, NMAE* and *IA.*

Table 3
Comparison between the different models on specific dates in April 2013. The best performances are marked in boldface.

Date	1-Apr	2-Apr	3-Apr	4-Apr	5-Apr	6-Apr	7-Apr	Average
					RMSE			
BPNN	0.1202	0.1149	0.1442	0.2333	0.1829	0.1596	0.1711	0.1656
ELMAN	0.1297	0.1235	0.1583	0.2400	0.1889	0.1691	0.1811	0.1745
GRNN	0.1174	0.1131	**0.1409**	**0.2306**	0.1838	0.1538	0.1650	0.1626
LLNF	0.1228	0.1160	0.1459	0.2328	0.1843	0.1562	0.1704	0.1658
DSNP-oLLNF	**0.0920**	**0.1002**	0.1485	0.2318	**0.1545**	**0.1272**	**0.1398**	**0.1420**
					sRMSE			
BPNN	0.0801	0.0766	0.0961	0.1555	0.1220	0.1064	0.1141	0.1104
ELMAN	0.0865	0.0823	0.1056	0.1600	0.1260	0.1128	0.1208	0.1163
GRNN	0.0783	0.0754	**0.0940**	**0.1537**	0.1225	0.1025	0.1100	0.1084
LLNF	0.0819	0.0773	0.0973	0.1552	0.1229	0.1041	0.1136	0.1105
DSNP-oLLNF	**0.0580**	**0.0634**	0.0990	0.1446	**0.0997**	**0.0814**	**0.0899**	**0.0909**
					NMAE			
BPNN	0.0642	0.0614	0.0780	0.1161	0.0984	0.0894	0.0924	0.0859
ELMAN	0.0685	0.0668	0.0863	0.1202	0.1040	0.0935	0.0956	0.0910
GRNN	0.0617	0.0596	**0.0760**	**0.1137**	0.0972	0.0833	0.0879	0.0830
LLNF	0.0653	0.0617	0.0810	0.1147	0.0994	0.0861	0.0914	0.0860
DSNP-oLLNF	**0.0443**	**0.0451**	0.0809	0.1156	**0.0805**	**0.0671**	**0.0720**	**0.0722**
					IA			
BPNN	0.9678	0.9694	0.9784	0.9472	0.9071	0.9718	0.9625	0.9609
ELMAN	0.9626	0.9654	0.9743	0.9432	0.8988	0.9688	0.9588	0.9569
GRNN	0.9705	0.9714	**0.9800**	**0.9499**	0.9112	0.9752	0.9666	0.9637
LLNF	0.9667	0.9695	0.9779	0.9460	0.9059	0.9735	0.9637	0.9610
DSNP-oLLNF	**0.9871**	**0.9816**	0.9768	0.9428	**0.9360**	**0.9888**	**0.9805**	**0.9705**

However, conclusions should not be reached based solely on experiments covering one month. Considering that wind power may fluctuate seasonally and over short periods, the experimental period is expanded to an entire year, from April 2013 to March 2014. The specific results of the larger experimental horizon are provided in Table 4. For instance, the data in the third row beside the "BPNN model" are the value of *RMSE*, which illustrates the accuracy of the prediction in the first week of May 2013, and there are 672 predicted values. The *RMSE*, *sRMSE*, *NMAE* and *IA* values of the DSNP-oLLNF model are 0.0864, 0.0542, 0.0406 and 0.9831, respectively; each value is the best compared with the control group models. Moreover, the merits of the hybrid model exist in all subjects of this experiment, not only one or several periods. Although there are some cases in which the superiority of the DSNP-oLLNF model is not notable in some error indicators, such as in June, July and October 2013, we can still confirm the improvements in the hybrid model in terms of the forecasting accuracy, as there are many uncertainties and influences of external factors. The improvements in these error indicators are notable.

The comparison of the forecast results is also illustrated in Fig. 5, from which we

can further confirm the merits of the developed hybrid DSNP-oLLNF model. For the Fig. 5, the part 1-4 are the forecasting performance of different models from May 1th to May 31th, 2013. To be specifically, part 1 is the forecasting values of four models, namely the hybrid DSNP-oLLNF model, BPNN model, ELMAN model and GRNN model, and the actual wind power values. It should be noted that the unit of wind power is MW (1 MW=1000 KW), which makes the difference of predicting curves look tiny. The part 2 is the Box-plot of the Mean Absolute Percentage Error for this four proposed contrast model as well as the statistical model ARIMA. It is visible that the performance of the hybrid DSNP-oLLNF is better than other four models. Besides, part 3 clearly depicts that the test values compared with the predicted values, from which we can obviously find that the forecasting values of the proposed DSNP-oLLNF method is very approximate to the test values. Moreover, as the evaluation of the whole performance, part 4 indicates that three error criteria of DSNP-oLLNF are the smallest between all methods.

In this paper, several processing strategies, including the DWT, SSA, NNCT and PSR are utilized in combination. To verify the necessity of the integration, a sequence of parallel comparisons are conducted between the hybrid DSNP-oLLNF model and the original LLNF model with single process strategy, including the DSN-OLLNF method and DP-OLLNF method, SP-OLLNF and P-OLLNF model. Just as the name implies, DSN-OLLNF is the model which contains DWT, SSA and NNCT model without the PSR, which means the form of mapping equation is $x_{t+5} = f\left(x_t, x_{t+1}, \ldots, x_{t+4}\right)$, as the faculty value of m is 5 and τ is 1. Besides, DP-OLLNF denotes the model without the SSA processing and SP-OLLNF without the DWT processing. As to the P-OLLNF model, there is only the PSR method included. The reason to conduct these comparisons is to verify whether the integration of processing strategy is necessary or if the integration is effective.

Table 4

Forecasting results obtained from different models.

Date	Apr. 2013	May. 2013	Jun. 2013	Jul. 2013	Aug. 2013	Sep. 2013	Oct. 2013	Nov. 2013	Dec. 2013	Jan. 2014	Feb. 2014	Mar. 2014	Average
							RMSE						
BPNN	0.1025	0.0602	0.0501	0.0557	0.0583	0.0599	0.0595	0.0572	0.0761	0.0774	0.0611	0.0480	0.0638
ELMAN	0.1053	0.0626	0.0515	0.0593	0.0630	0.0617	0.0613	0.0600	0.0772	0.0802	0.0669	0.0506	0.0666
GRNN	0.1052	0.0671	0.0530	0.0794	0.0666	0.0693	0.0607	0.0606	0.0793	0.0945	0.0713	0.0573	0.0720
LLNF	0.1010	0.0602	0.0492	0.0555	0.0580	0.0613	0.0593	0.0580	0.0760	0.0724	0.0610	0.0593	0.0643
DSNP-oLLNF	**0.0864**	**0.0489**	**0.0385**	**0.0448**	**0.0471**	**0.0474**	**0.0485**	**0.0461**	**0.0629**	**0.0546**	**0.0498**	**0.0368**	**0.0510**
							sRMSE						
BPNN	0.0683	0.0402	0.0334	0.0371	0.0389	0.0399	0.0397	0.0381	0.0507	0.0516	0.0407	0.0320	0.0426
ELMAN	0.0702	0.0417	0.0343	0.0396	0.0420	0.0411	0.0409	0.0400	0.0515	0.0535	0.0446	0.0338	0.0444
GRNN	0.0701	0.0447	0.0353	0.0530	0.0444	0.0462	0.0405	0.0404	0.0528	0.0630	0.0476	0.0382	0.0480
LLNF	0.0673	0.0401	0.0328	0.0370	0.0386	0.0409	0.0395	0.0386	0.0507	0.0483	0.0407	0.0317	0.0422
DSNP-oLLNF	**0.0542**	**0.0293**	**0.0223**	**0.0265**	**0.0281**	**0.0283**	**0.0290**	**0.0274**	**0.0386**	**0.0330**	**0.0299**	**0.0212**	**0.0306**
							NMAE						
BPNN	0.0535	0.0309	0.0253	0.0282	0.0297	0.0306	0.0297	0.0279	0.0353	0.0363	0.0307	0.0240	0.0318
ELMAN	0.0559	0.0320	0.0260	0.0301	0.0322	0.0309	0.0314	0.0291	0.0364	0.0388	0.0337	0.0255	0.0335
GRNN	0.0543	0.0352	0.0266	0.0413	0.0341	0.0362	0.0309	0.0294	0.0379	0.0427	0.0370	0.0297	0.0363
LLNF	0.0527	0.0306	0.0245	0.0281	0.0294	0.0314	0.0299	0.0281	0.0359	0.0347	0.0306	0.0236	0.0316
DSNP-oLLNF	**0.0406**	**0.0195**	**0.0139**	**0.0176**	**0.0189**	**0.0186**	**0.0194**	**0.0175**	**0.0244**	**0.0214**	**0.0198**	**0.0194**	**0.0209**
							IA						
BPNN	0.9699	0.9893	0.9866	0.9891	0.9915	0.9902	0.9794	0.9538	0.9567	0.9700	0.9833	0.9907	0.9792
ELMAN	0.9671	0.9885	0.9863	0.9878	0.9901	0.9899	0.978	0.9496	0.9551	0.9674	0.9797	0.9896	0.9774
GRNN	0.9668	0.9865	0.9854	0.9789	0.9887	0.9865	0.9782	0.9457	0.9507	0.9542	0.9752	0.9864	0.9736
LLNF	0.9701	0.9894	0.9876	0.9892	0.9916	0.9896	0.9797	0.9523	0.9563	0.9745	0.9834	0.9910	0.9795
DSNP-oLLNF	**0.9831**	**0.9999**	**0.9980**	**0.9994**	**0.9919**	**0.9991**	**0.9902**	**0.9655**	**0.9706**	**0.9908**	**0.9942**	**0.9953**	**0.9898**

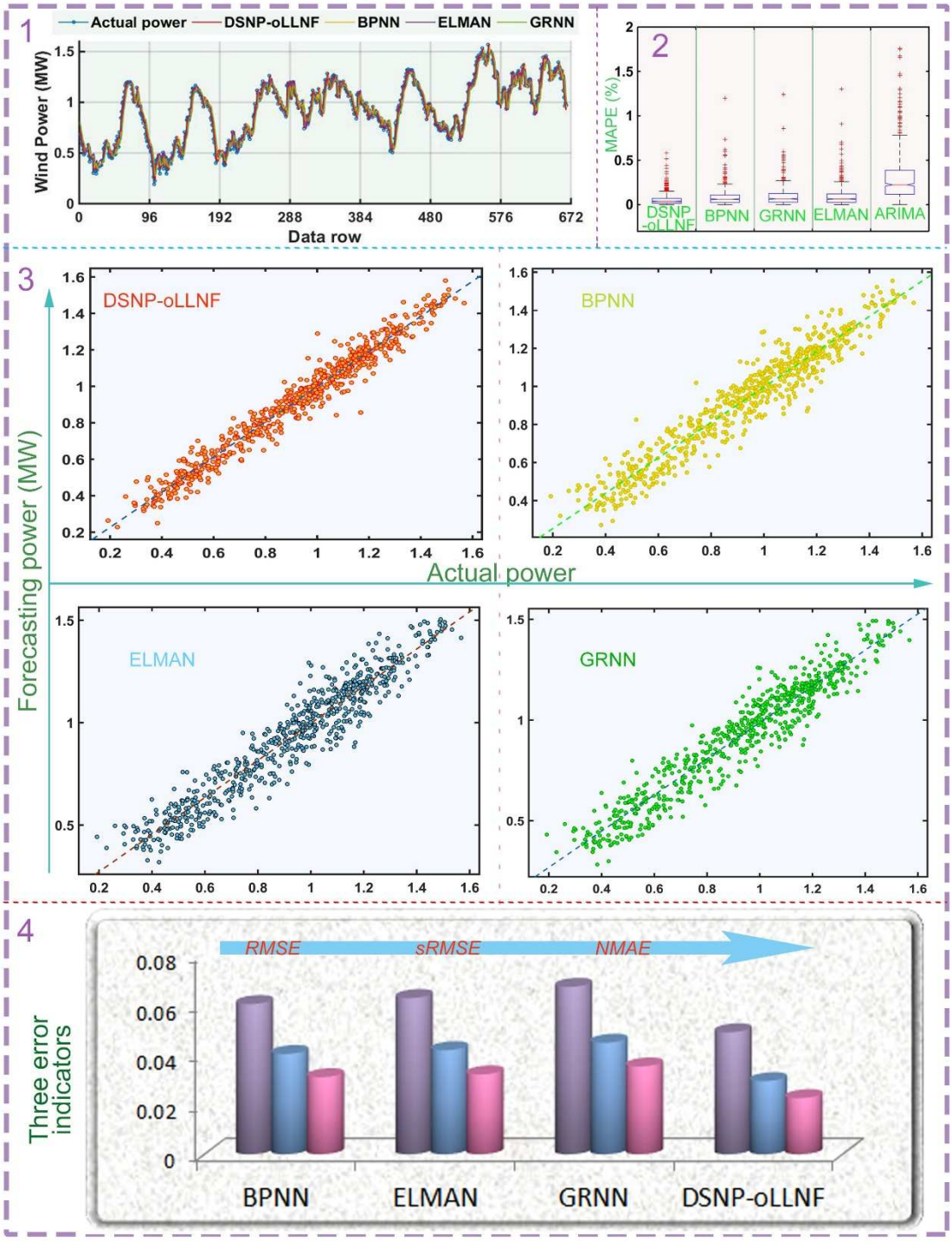


Fig. 5. Comparison of forecasting results, May 2013, Aksay region.

The results of the comparison are listed in Table 5. The experimental period also spans an entire year, covering twelve months from April 2013 to March 2014. The longer study period is considered to reduce the accidental error, including the seasonal factors and periodical influences, so as to reveal the actual rules. The experimental structure is similar to that in the former experiment, except for the experimental models. According to the results shown in Table 5, the single process strategy is inferior to the hybrid DSNP-oLLNF model, and this phenomenon is observed in all four error indicators. For instance, in September 2013, the hybrid model leads to reductions of 22.6754%, 39.3068%, 18.1347% and 17.9931% in the *RMSE*,

reductions of 30.8068%, 45.6814%, 26.6839% and 26.6839% in the *sRMSE*, and reductions of 40.5751%, 53.2663%, 36.3014% and 36.5188% in the *NMAE* compared with the DSN-OLLNF, DP-OLLNF, SP-OLLNF and P-OLLNF models, respectively. Additionally, the *IA* of the hybrid model is 0.65%, 1.25%, 0.44% and 0.44% higher than those of the DSN-OLLNF and DP-OLLNF, SP-OLLNF and P-OLLNF models, respectively. The phenomenon also exists in other subjects. Thus, these results reflect that the hybrid model proposed in this paper can outperform other models based on the *RMSE*, *sRMSE* and *NMAE* values. The effectiveness and validity of hybrid model, as indicated by the *IA* values, also reveal the superiority and stability of the DSNP-oLLNF model for wind power forecasting.

Table 5
Performance comparison between the single processing methods and the hybrid model.

Date	Apr, 2013	May, 2013	Jun, 2013	Jul, 2013	Aug, 2013	Sep, 2013
			DSN-OLLNF			
RMSE	0.1013	0.0599	0.0495	0.0559	0.058	0.0613
sRMSE	0.0675	0.0399	0.033	0.0373	0.0387	0.0409
NMAE	0.053	0.0304	0.0245	0.0284	0.0293	0.0313
IA	0.97	0.9895	0.9874	0.989	0.9916	0.9896
			DP-OLLNF			
RMSE	0.1116	0.075	0.0683	0.0608	0.0612	0.0781
sRMSE	0.0744	0.05	0.0456	0.0405	0.0408	0.0521
NMAE	0.0585	0.0387	0.0356	0.0305	0.031	0.0398
IA	0.9651	0.9838	0.9765	0.9871	0.9908	0.9836
			SP-OLLNF			
RMSE	0.165	0.1004	0.0594	0.0486	0.0556	0.0579
sRMSE	0.11	0.067	0.0396	0.0324	0.037	0.0386
NMAE	0.0854	0.0528	0.0301	0.024	0.0282	0.0292
IA	0.9617	0.97	0.9896	0.9879	0.9891	0.9917
			P-OLLNF			
RMSE	0.1681	0.1016	0.05982	0.0488	0.0553	0.0578
sRMSE	0.1121	0.0678	0.0399	0.0325	0.0369	0.0386
NMAE	0.087	0.0531	0.0304	0.0243	0.0281	0.0293
IA	0.9601	0.9695	0.9895	0.9877	0.9892	0.9917
			DSNP-oLLNF			
RMSE	0.0864	0.0489	0.0385	0.0448	0.0471	0.0474
sRMSE	0.0542	0.0293	0.0223	0.0265	0.0281	0.0283
NMAE	0.0406	0.0195	0.0139	0.0177	0.0189	0.0186
IA	0.9831	0.9999	0.998	0.9994	0.9959	0.9961

Table 5
(Continue).

Date	Oct, 2013	Nov, 2013	Dec, 2013	Jan, 2014	Feb, 2014	Mar, 2014
			DSN-OLLNF			
RMSE	0.0596	0.0581	0.0759	0.0684	0.061	0.048
sRMSE	0.0397	0.0387	0.0506	0.0456	0.0407	0.032
NMAE	0.0299	0.0285	0.0361	0.0329	0.0305	0.0239
IA	0.9796	0.9521	0.9566	0.9777	0.9834	0.9908
			DP-OLLNF			
RMSE	0.0654	0.0746	0.0998	0.0821	0.0799	0.0577

sRMSE	0.0436	0.0498	0.0665	0.0547	0.0533	0.0385
NMAE	0.0329	0.0369	0.0483	0.041	0.0413	0.0294
IA	0.9757	0.9263	0.9301	0.9696	0.9725	0.9869
			SP-OLLNF			
RMSE	0.0603	0.0591	0.058	0.0744	0.0671	0.061
sRMSE	0.0402	0.0394	0.0387	0.0496	0.0448	0.0403
NMAE	0.0309	0.0298	0.0283	0.0353	0.0326	0.0303
IA	0.99	0.9798	0.9519	0.9585	0.9787	0.9838
			P-OLLNF			
RMSE	0.5999	0.0591	0.0571	0.0755	0.0663	0.0603
sRMSE	0.04	0.0394	0.038	0.0503	0.0442	0.0402
NMAE	0.0307	0.0297	0.0279	0.0355	0.0321	0.0301
IA	0.9901	0.9799	0.9536	0.9572	0.9795	0.9839
			DSNP-oLLNF			
RMSE	0.0485	0.0461	0.0629	0.0546	0.0498	0.0268
sRMSE	0.019	0.0274	0.0386	0.033	0.0299	0.0212
NMAE	0.0194	0.0175	0.0244	0.0214	0.0198	0.0134
IA	0.9902	0.9855	0.9706	0.9908	0.9942	0.9963

Furthermore, the proposed hybrid model is compared with a sequence of models presented in the existing literature, and the *RMSE*, *sRMSE* and *NMAE* are summarized in Table 6. Because of the generality of the *sRMSE*, the ratio of the *sRMSE* of the proposed model to the *sRMSE* of the compared models is also shown in Table 6. A ratio of *sRMSE* > 1 indicates that the proposed DSNP-oLLNF model is worse than the predictions of the other methods. By contrast, a ratio of *sRMSE* < 1 indicates that the DSNP-oLLNF model outperforms the competing prediction models. Thus, smaller values for the *sRMSE* ratio indicates a better performance of the proposed model. The comparison results listed in Table 6 reveal the merits of the proposed hybrid model.

Table 6
Wind power forecasting results of 2013-7 in Aksay.

Method	**Error indicators**			
	RMSE	***sRMSE***	***NMAE***	***sRMSE ratio***
ARIMA	0.0702	0.0468	0.0343	0.6667
ANFIS [36]	0.0566	0.0377	0.0287	0.8276
RBF [37]	0.056	0.0374	0.0277	0.8342
Khan et al.'s model [38]	0.0560	0.0373	0.0283	0.8365
Sadiq & Sultana's model [39]	0.0558	0.0372	0.0282	0.8387
Da & Xiurun's model [40]	0.0553	0.0368	0.0278	0.8478
Shi et al.'s model [41]	0.0551	0.0367	0.0272	0.8501
ELM [42]	0.055	0.0366	0.0278	0.8525
Nieto et al.'s model [43]	0.0545	0.0363	0.0272	0.8595
Lesniewski & Bartoszewicz's model [44]	0.0550	0.036	0.0281	0.8667
DSNP-oLLNF	**0.0367**	**0.0212**	**0.0133**	-

Finally, to determine whether precision difference between the hybrid model and the existing models is statistically significant, we consider the loss functions as forecasting error evaluation functions to conduct the *DM* test directly between the hybrid proposed DSNP-oLLNF model and the traditional LLNF, BPNN, ELMAN and GRNN models. The results for the *DM* test are presented in Table 7. All of these indexes are significant at a level of at least 10%. Thus, we can draw the conclusion that the forecasting accuracy of the DSNP-oLLNF model is significantly different

from the contrast models, i.e., the developed hybrid approach is superior and reliable.

Table 7
Results of the *DM* test of Aksay.

DM	DSNP-oLLNF			
	April	July	October	January
LLNF	2.9552**	2.2511**	1.7883**	3.4725*
BPNN	1.0015	2.6903*	1.8054**	4.698*
ELMAN	3.5116*	4.0864*	4.4833*	6.3507*
GRNN	2.6122*	1.6205***	4.2671*	3.6149*

*significant at the 1% level.
**significant at the 5% level.
***significant at the 10% level.

5.2. Experiment II

In this experiment, the data are collected from the Yumen-Guazhou wind farm, which is located between Yumen city and Guazhou city of Gansu, China. This wind farm was built in 2009 with an installed capacity of 3,000 MW, and the capacity has been increasing as the government boosts the investment. The typical turbines utilized in this wind farm are purchased from Huarui Company (Beijing, China), model SL1500/D82. The rated power is 1,500 kW, and the height of the hub is 70 m. The characteristics of this region are also shown in Table 2.

In contrast to experiment I, the performance of the proposed model is demonstrated by multistep prediction in this part. For selected step of prediction horizons λ, there have 1 step = 15 min, 2 steps = 30 min, 4 steps = 1 h, 8 steps = 2 h. It is worth noting that multistep forecasting model belongs to the category of Multi-Input-Multi-Output (MIMO) strategy, which can be expressed as a multi output equation:

$$\left[x_{t+1+(m-1)\tau}, x_{t+2+(m-1)\tau}, \ldots, x_{t+\lambda+(m-1)\tau}\right] = f'\left(x_t, x_{t+\tau}, \ldots, x_{t+(m-1)\tau}\right) \tag{18}$$

where λ is the step of forecast horizon, $f'(\cdot)$ is a mapping function and $x_t(t=1,2,\ldots,N)$ denotes the processed wind power series. In experiment II, as shown in Fig. 6, the values of the delay time τ and the embedding dimension m are 3 and 5, respectively.

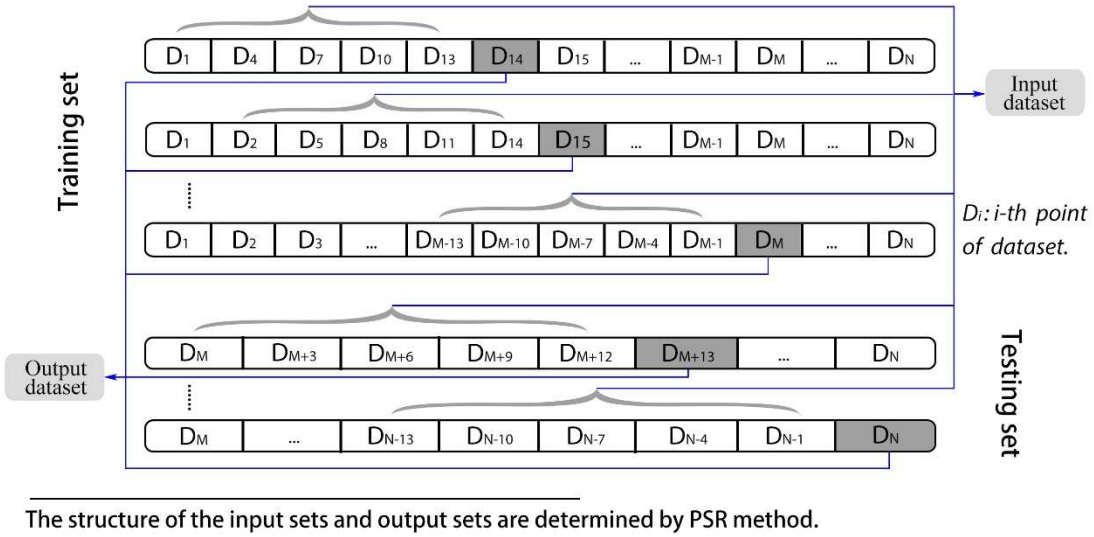


Fig. 6. Form of the dataset adopted in experiment II.

In experiment II, wind power series span a horizon from May 1, 2013 to July 10,

2014 with a time interval of **15 minutes**. Totally, there are 38682 numbers of wind power data collected, and the fraction of the missing data is 3.82%, specifically, the number of entry/missing data is 38682/1477. Fig. 3 shows the pictures of 15 min average wind power from Yumen region along with the time variation.

When conducting experiment II, power data covering a whole month are set as training sets. For instance, the processed wind power series from June 1 to June 31, 2013 are used to train the model, then the trained model is utilized to predict the following wind power series values of the first week in April. Different from experiment I, all the results of experiment II are obtained by multistep forecasting. The reason to choose this form is to testify the generalization of proposed hybrid model, meanwhile, to guarantee the requirements of longer term prediction in practise, e.g. 2 h. In experiment II, the maximum step of horizon is 8 since the hybrid DSNP-oLLNF model is proposed for short term wind power forecasting originally.

Table 8 presents the *RMSE*, *sRMSE* and *NMAE* results for selected prediction horizons, from which we can find that in most cases, the DSNP-oLLNF model provides the best forecasting accuracy. Moreover, the differences between the proposed hybrid model and the compared models are significant. Table 8 also reveals the asymmetry: the prediction accuracy decreases with the increasing in the study horizon. The *RMSE*, *sRMSE* and *NMAE* shown in Table 8, including all forecasting horizons, demonstrate the accuracy and reliability of the proposed hybrid model. The detail results of experiment II are also dipicted in Fig. 7.

Similarly, a comparison between the proposed DSNP-oLLNF model and a sequence of models presented in the existing literature, in terms of the *RMSE*, *sRMSE*, *NMAE* and *sRMSE* ratio are performed, which are shown in Table 9. The comparison reveals the superiority of the hybrid approach compared with the other methods that have been proposed in the literature. Moreover, the ratio of the *sRMSE* of the proposed hybrid model to the *sRMSE* of the compared models is provided in Table 9, which indicates the improvement in the prediction performance. For instance, for the ARIMA, ANFIS and RBF models, the values of the *sRMSE* ratio are 0.5169, 0.8239 and 0.8389, respectively.

In addition, the results of the *DM* forecasting accuracy tests, which are listed in Table 10, indicate the superiority of the proposed hybrid model, as the calculated *DM* statistic of the DSNP-oLLNF model are significant at different levels in most cases. Therefore, there are notable differences in forecasting accuracy between the proposed hybrid model and the existing models. As indicated by the comparison results in Tables 8 and 9 and demonstrated by the *DM* statistics, the developed hybrid model has superior performance over other approaches.

In summary, the developed DSNP-oLLNF model is an option for short term wind power forecasting and can be taken into consideration on daily scheduling and management. In a nutshell, the satisfactory performance can be explained that the proposed integrated processing strategies, including the DWT, SSA, NNCT and PSR, can not only extract the traits of the original wind power series but also find the most proper structure of input sets. Besides, the SOA as an effective technique, which is able to optimize the initial parameters of LLNF, can improve the forecasting accuracy effectively.

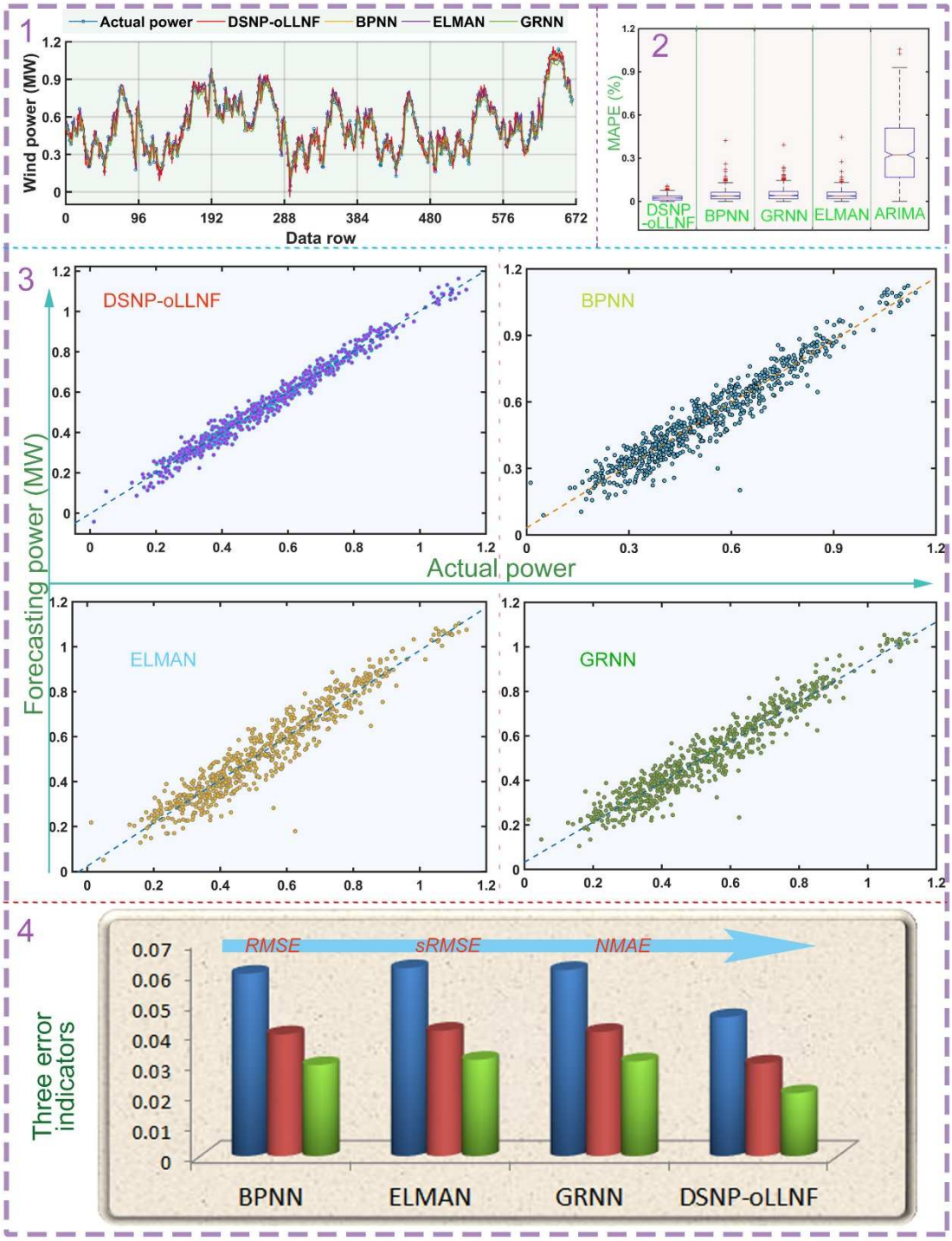


Fig. 7. Comparison of forecasting results, October 2013, Yumen region.

680 **Table 8**
681 Error indexes, time frame of 5/1/13 to 7/10/14, all time horizons.

Step (s)	Date	BPNN			ELMAN			GRNN			LLNF			DSNP-oLLNF		
		RMSE	***sRMSE***	***NMAE***	***RMSE***	***sRMSE***	***NMAE***	***RMSE***	***sRMSE***	***NMAE***	***RMSE***	***sRMSE***	***NMAE***	***RMSE***	***sRMSE***	***NMAE***
1	**Apr, 2013**	0.103	0.0687	0.0537	0.1058	0.0705	0.0556	0.1074	0.0716	0.0554	0.0989	0.0659	0.0515	0.0861	0.0541	0.0403
	Jul, 2013	0.0566	0.0377	0.0288	0.0577	0.0385	0.0291	0.0603	0.0402	0.0312	0.0557	0.0371	0.0283	0.0449	0.0266	0.0176
	Oct, 2013	0.0594	0.0396	0.0299	0.0633	0.0422	0.0318	0.0608	0.0405	0.031	0.0593	0.0395	0.0299	0.0486	0.029	0.0194
	Jan, 2014	0.0955	0.0637	0.0413	0.0928	0.0618	0.0437	0.0913	0.0609	0.0411	0.0696	0.0464	0.0336	0.0547	0.0332	0.0214
2	**Apr, 2013**	0.1913	0.1275	0.0995	0.1891	0.1261	0.0992	0.1902	0.1268	0.1005	0.1765	0.1177	0.0916	0.1338	0.0859	0.0649
	Jul, 2013	0.1053	0.0702	0.0537	0.1104	0.0736	0.0568	0.1077	0.0718	0.0548	0.106	0.0707	0.0542	0.08	0.05	0.0361
	Oct, 2013	0.1149	0.0766	0.0609	0.1154	0.0769	0.0621	0.1139	0.0759	0.0595	0.1151	0.0768	0.0613	0.0859	0.0539	0.0397
	Jan, 2014	0.1645	0.1097	0.0845	0.1542	0.1028	0.0779	0.2097	0.1398	0.0969	0.1485	0.099	0.0751	0.1027	0.0651	0.0461
4	**Apr, 2013**	0.2644	0.1763	0.1402	0.2787	0.1858	0.1495	0.2797	0.1865	0.1501	0.2752	0.1835	0.1477	0.1703	0.1102	0.0866
	Jul, 2013	0.144	0.096	0.0766	0.1515	0.101	0.0809	0.1564	0.1042	0.0806	0.1552	0.1035	0.082	0.1072	0.0681	0.0511
	Oct, 2013	0.1464	0.0976	0.0769	0.153	0.102	0.0828	0.1722	0.1148	0.0894	0.1587	0.1058	0.0868	0.11	0.07	0.0536
	Jan, 2014	0.2071	0.1381	0.1092	0.2049	0.1366	0.1094	0.2503	0.1699	0.6086	0.2274	0.1516	0.1183	0.1369	0.0879	0.0659
8	**Apr, 2013**	0.3225	0.215	0.1716	0.3151	0.2101	0.1635	0.3329	0.2219	0.1793	0.3299	0.22	0.1762	0.2268	0.1478	0.1171
	Jul, 2013	0.1775	0.1184	0.096	0.1828	0.1219	0.1005	0.1965	0.131	0.1034	0.1904	0.1269	0.1019	0.1286	0.0824	0.0639
	Oct, 2013	0.2042	0.1361	0.6336	0.1978	0.1318	0.1071	0.208	0.1387	0.1089	0.1983	0.1322	0.1059	0.1337	0.0858	0.0662
	Jan, 2014	0.2595	0.173	0.1404	0.2392	0.1595	0.1315	0.2697	0.1798	0.144	0.2627	0.1751	0.1432	0.181	0.1173	0.0922

Table 9
Wind power forecasting results for 07/2013 in Yumen.

Method	Error indicators			
	RMSE	***sRMSE***	***NMAE***	***sRMSE ratio***
ARIMA	0.0801	0.0534	0.0394	0.5169
ANFIS [36]	0.0503	0.0335	0.0249	0.8239
RBF [37]	0.0494	0.0329	0.0245	0.8389
Khan et al.'s model [38]	0.0492	0.0328	0.0244	0.8415
Sadiq & Sultana's model [39]	0.0488	0.0325	0.0243	0.8456
Da & Xiurun's model [40]	0.0486	0.0323	0.0242	0.8476
Shi et al.'s model [41]	0.0499	0.0333	0.0245	0.8288
ELM [42]	0.0494	0.0329	0.0247	0.8389
Nieto et al.'s model [43]	0.0501	0.0334	0.0247	0.8263
Lesniewski & Bartoszewicz's model [44]	0.0488	0.0326	0.0242	0.8466
DSNP-oLLNF	**0.0289**	**0.0176**	**0.0119**	-

Table 10
Results of the *DM* test in Yumen.

DM	**DSNP-oLLNF**			
	March	**June**	**September**	**February**
LLNF	2.7984*	1.9195**	1.2831***	1.8546**
BPNN	2.8321*	2.6681*	1.9036**	1.6313***
ELMAN	4.6446*	12.019*	4.7201*	3.2956*
GRNN	2.0956**	3.7672*	6.3848*	3.3167*

*significant at the 1% level.
**significant at the 5% level.
***significant at the 10% level.

6. Conclusions

Affected by various environmental and meteorological factors, wind power series usually present nonlinearity, non-stationary and chaos, making it a tough question to find the internal tendency and make accurate forecasting. In this paper, DWT method is adopted to eliminate the stochastic volatility or the noise signals from raw wind power series while SSA approach is used to extract the trend and periodic from original series, then NNCT theory together with the CS algorithm are introduced to combine these two processing approaches retaining strengths of each method. Next, PSR method is adopted to select the most proper parallel structure of the input sets and output sets, before we present the optimized local linear fuzzy neural network as the predictor to make short term wind power forecasting. The contribution percentages of the processing strategy and the SOA optimization algorithm to promote the single LLNF model are investigated based on two experiments. In experiment I, from April, 2013 to March, 2014, the average forecasting results of the DSNP-oLLNF model are 0.051, 0.0306, 0.0209 and 0.9898, respectively, with respect to *RMSE*, *sRMSE*, *NMAE* and *IA*. Compared to the best performances obtained by the conventional models (0.0638, 0.0422, 0.0316 and 0.9795, respectively, with respect to *RMSE*, *sRMSE*, *NMAE* and *IA*), the forecasting accuracy improvements of the DSNP-oLLNF model are 20.0627%, 27.4882%, 33.8608% and 1.0516%, respectively. On the other hand, the superiority of the hybrid DSNP-oLLNF model also exist in multistep forecasting of wind power, which are conducted in experiment II.

In summary, the results of these two experiments show that: (1) the developed hybrid processing strategy can improve the prediction performance considerably, which is discovered that the combination of single SSA and DWT by NNCT method is feasible; (2) the hybrid DSNP-oLLNF model outperforms conventional single models in short term wind power forecasting, which indicates that the integration of

processing strategy, SOA algorithm and single LLNF is a right way to promote the capacity of the single LLNF network in short term wind power prediction and (3) the developed DSNP-oLLNF model also precedes these hybrid models proposed in recent researches with specific comparisons shown in Tables 6 and 9.

Since wind energy is the world's fastest growing source of clean and renewable energy, the predictability of wind power generation is important for the safety of the wind power conversion and integration of wind energy into the power system. Improving forecasting accuracy is an essential yet often difficult task, which the relevant decision makers constantly have to face. To be honest, despite the effectiveness of using the hybrid DSNP-oLLNF model illustrated above, there certainly exist some limitations. For instance, only wind power series with time interval of 15 minutes are used to evaluate the performances of the developed model in this paper. A possible topic of further research is to make forecasting with the monthly or even quarterly data.

Conflict of interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant No. 71573034).

Appendix: List of abbreviations

Here, all of the terms used in this paper and their definitions are listed in alphabetical order:

Nomenclature			
ANNs	Artificial neural networks	***MAE***	Mean absolute error
ARIMA	Autoregressive integrated moving average model	***MAPE***	Mean absolute percentage error
ARMA	Autoregressive moving average process	**MIMO**	Multi-Input-Multi-Output
BPNN	Back propagation neural network	***MSE***	Mean square error
CRO	Coral reefs optimization	***NMAE***	Normalized mean-absolute of errors
CS	Cuckoo search algorithm	**NNCT**	The no negative constraint combination strategy
DM	Diebold–Mariano test	**NWP**	Numerical weather prediction
DP-oLLNF	Hybrid model without SSA	**P-oLLNF**	Hybrid model without SSA and DWT
DSN-oLLNF	Hybrid model without PSR	**PRW**	Persistent random walk
DSNP-oLLNF	Proposed integrated model	**PSR**	Phase space reconstruction
DWT	Discrete wavelet transform	**RBF**	Radical basis function network
ELMs	Extreme learning machine	***RMSE***	Root-mean-squares of errors
ELMAN	Elman network	**SAM**	Seasonal adjustment method
EMD	Empirical mode decomposition	**SOA**	Seeker optimization algorithm
GRNN	Generalized regression neural network	**SOM**	Self-organization maps
HS	Harmony search	**SP-oLLNF**	Hybrid model without DWT
IA	Index of agreement	***sRMSE***	Standardized root-mean-squares of errors
LLM	Local linear model	**SSA**	Singular spectrum analysis
LLNF	Local linear neuro-fuzzy model	**SVD**	Singular value decomposition
LoLiMoT	Local linear modelling tree	**SVMs**	Support vector machines
		WRF	Weather research and

algorithm	forecasting

References

[1] B.-A. Schuelke-Leech, B. Barry, M. Muratori, B.J. Yurkovich, Big Data issues and opportunities for electric utilities, Renew. Sustain. Energy Rev. 52 (2015) 937–947.

[2] C. Croonenbroeck, G. Stadtmann, Minimizing asymmetric loss in medium-term wind power forecasting, Renew. Energy. 81 (2015) 197–208.

[3] Y. Feng, H. Lin, S.L. Ho, J. Yan, J. Dong, S. Fang, Y. Huang, Overview of wind power generation in China: Status and development, Renew. Sustain. Energy Rev. 50 (2015) 847–858.

[4] C. Stathopoulos, A. Kaperoni, G. Galanis, G. Kallos, Wind power prediction based on numerical and statistical models, J. Wind Eng. Ind. Aerodyn. 112 (2013) 25–38.

[5] E.G. Salcedo-Sanz S, Ángel M. Pérez-Bellido, Ortiz-García E G, A. Portilla-Figueras, L. Prieto, D. Paredes, Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction, Renew. Energy. 34 (2009) 1451–1457.

[6] S. Salcedo-Sanz, E.G. Ortiz-García, A.M. Pérez-Bellido, E. Portilla-Figueras, L. Prieto, D. Paredes, F. Correoso, Performance comparison of Multilayer Perceptrons and Support vector Machines in a Short-term Wind speed Prediction Problem, Neural Netw. World. 19 (2009) 37–51.

[7] F. Cassola, M. Burlando, Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output, Appl. Energy. 99 (2012) 154–166.

[8] J. Zhao, Z.H. Guo, Z.Y. Su, Z.Y. Zhao, X. Xiao, F. Liu, An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed, Appl. Energy. 162 (2016) 808–826.

[9] J. Hu, J. Wang, K. Ma, A hybrid technique for short-term wind speed prediction, Energy. 81 (2015) 563–574.

[10] A.M. Foley, P.G. Leahy, A. Marvuglia, E.J. McKeogh, Current methods and advances in forecasting of wind power generation, Renew. Energy. 37 (2012) 1–8.

[11] S.A. Pourmousavi Kani, M.M. Ardehali, Very short-term wind speed prediction: A new artificial neural network-Markov chain model, Energy Convers. Manag. 52 (2011) 738–745.

[12] P. Pinson, C. Chevallier, G.N. Kariniotakis, Trading wind generation from short-term probabilistic forecasts of wind power, IEEE Trans. Power Syst. 22 (2007) 1148–1156.

[13] W. Zhang, J. Wang, J. Wang, Z. Zhao, M. Tian, Short-term wind speed forecasting based on a hybrid model, Appl. Soft Comput. J. 13 (2013) 3225–3233.

[14] M. Khashei, M. Bijari, G.A. Raissi Ardali, Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and Artificial Neural Networks (ANNs), Neurocomputing. 72 (2009) 956–967.

[15] G.H. Riahy, M. Abedi, Short term wind speed forecasting for wind turbine applications using linear prediction method, Renew. Energy. 33 (2008) 35–41.

[16] J.L. Torres, A. García, M. De Blas, A. De Francisco, Forecast of hourly average wind speed with ARMA models in Navarre (Spain), Sol. Energy. 79 (2005) 65–77.

[17] S. Salcedo-Sanz, A. Pastor-Sánchez, L. Prieto, A. Blanco-Aguilera, R. García-Herrera, Feature selection in wind speed prediction systems based on a hybrid coral reefs optimization - Extreme learning machine approach, Energy Convers. Manag. 87 (2014) 10–18.

[18] S. Salcedo-Sanz, E.G. Ortiz-García, Á.M. Pérez-Bellido, A. Portilla-Figueras, L. Prieto, Short term wind speed prediction based on evolutionary support vector regression algorithms, Expert Syst. Appl. 38 (2011) 4052–4057.

[19] Y. Jiang, Z. Song, A. Kusiak, Very short-term wind speed forecasting with Bayesian structural break model, Renew. Energy. 50 (2013) 637–647.

[20] G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting, Appl. Energy. 87 (2010) 2313–2320.

[21] E.G. Ortiz-García, S. Salcedo-Sanz, Á.M. Pérez-Bellido, J. Gascón-Moreno, J.A. Portilla-Figueras, L. Prieto, Short-term wind speed prediction in wind farms based on banks of support vector machines, Wind Energy. 14 (2011) 193–207.

[22] S. Salcedo-Sanz, A. Pastor-Sánchez, J. Del Ser, L. Prieto, Z.W. Geem, A Coral Reefs Optimization algorithm with Harmony Search operators for accurate wind speed prediction, Renew. Energy. 75 (2015) 93–101.

[23] H. Liu, C. Chen, H. Tian, Y. Li, A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks, Renew. Energy. 48 (2012) 545–556.

[24] K. Gnana Sheela, S.N. Deepa, Neural network based hybrid computing model for wind speed prediction, Neurocomputing. 122 (2013) 425–429.

[25] H. Liu, H.Q. Tian, X.F. Liang, Y.F. Li, Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks, Appl. Energy. 157 (2015) 183–194.

[26] L. Xiao, J. Wang, Y. Dong, J. Wu, Combined forecasting models for wind energy forecasting: A case study in China, Renew. Sustain. Energy Rev. 44 (2015) 271–288.

[27] M. Abdollahzade, A. Miranian, H. Hassani, H. Iranmanesh, A new hybrid enhanced local linear neuro-fuzzy model based on the optimized singular spectrum analysis and its application for nonlinear and chaotic time series forecasting, Inf. Sci. (Ny). 295 (2015) 107–125.

[28] E. Carrizosa, A. V. Olivares-Nadal, P. Ramírez-Cobo, Time series interpolation via global optimization of moments fitting, Eur. J. Oper. Res. 230 (2013) 97–112.

[29] G.J. Osório, J.C.O. Matias, J.P.S. Catalão, Short-term wind power forecasting using adaptive neuro-fuzzy inference system combined with evolutionary particle swarm optimization, wavelet transform and mutual information, Renew. Energy. 75 (2015) 301–307.

[30] A. Tascikaraoglu, B.M. Sanandaji, K. Poolla, P. Varaiya, Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform, Appl. Energy. 165 (2016) 735–747.

[31] H. Hassani, A. Webster, E.S. Silva, S. Heravi, Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis, Tour. Manag. 46 (2015) 322–335.

[32] F. Takens, Dynamical systems and turbulence, in: Lecture notes in mathematics, Brownian Motion Its Appl. to Math. Anal. École d'Été Probab. Saint-Flour XLIII-2013. 341 (1981) viii,164.

[33] D.I.O. Nelles, D.I.O. Bänfer, J. Kainz, J. Beer, Local model networks, Atzelektronik Worldw. 3 (2008) 36–39.

[34] M.J. Navardi, B. Babaghorbani, A. Ketabi, Efficiency improvement and torque ripple minimization of Switched Reluctance Motor using FEM and Seeker Optimization Algorithm, Energy Convers. Manag. 78 (2014) 237–244.
[35] F.X. Diebold, R.S. Mariano, Comparing Predictive Accuracy, J. Bus. Econ. Stat. 13 (1995) 253–265.
[36] S. Amirkhani, S. Nasirivatan, a. B. Kasaeian, a. Hajinezhad, ANN and ANFIS models to predict the performance of solar chimney power plants, Renew. Energy. 83 (2015) 597–607.
[37] G. Melagraki, A. Afantitis, H. Sarimveis, O. Igglessi-Markopoulou, A. Alexandridis, A novel RBF neural network training methodology to predict toxicity to Vibrio fischeri, Mol. Divers. 10 (2006) 213–221.
[38] S. Khan, A.R. Baig, W. Shahzad, A novel ant colony optimization based single path hierarchical classification algorithm for predicting gene ontology, Appl. Soft Comput. 16 (2014) 34–49.
[39] M. Sadiq, S. Sultana, Computational Intelligence in Data Mining - Volume 1, Smart Innov. Syst. Technol. 31 (2015) 213–220.
[40] Y. Da, G. Xiurun, An improved PSO-based ANN with simulated annealing technique, Neurocomputing. 63 (2005) 527–533.
[41] Y. Shi, D.L. Yu, Y. Tian, Y. Shi, Air-fuel ratio prediction and NMPC for SI engines with modified Volterra model and RBF network, Eng. Appl. Artif. Intell. 45 (2015) 313–324.
[42] C. Wan, Z. Xu, P. Pinson, Z.Y. Dong, K.P. Wong, Probabilistic forecasting of wind power generation using extreme learning machine, IEEE Trans. Power Syst. 29 (2014) 1033–1044.
[43] P.J.G. Nieto, E. García-Gonzalo, J.R.A. Fernández, C.D. Muñiz, A hybrid PSO optimized SVM-based model for predicting a successful growth cycle of the Spirulina platensis from raceway experiments data, Ecol. Eng. 81 (2015) 534–542.
[44] P. Lesniewski, A. Bartoszewicz, Advances in Systems Science, Adv. Intell. Syst. Comput. 240 (2014) 45–55.

Highlights

- The hybrid forecasting approach can improve the accuracy of wind power forecasting significantly.
- A novel integrated processing strategy has been developed to filter and extract the raw datasets and then select the most proper parallel structure of input sets.
- Traditional local linear neural-fuzzy model is optimized by seeker optimization algorithm.
- The hybrid model outperforms than the traditional methods.